Bayesian inference of neural connectivity in a population of neurons from calcium imaging data

us

August 11, 2009

Contents

Abstract

Deducing the structure of neural circuits from neural activity is one of the central problems of modern neuroscience. Recent studies have examined this problem in the context of neural activity observations for dozens of neurons using multi-electrode techniques. Here, we present a model-based optimal approach for inferring structure of neural micro-circuits containing hundreds of neurons observed optically with calcium imaging. While calcium imaging allows less intrusive observations of neural activity in much larger

populations than multi-electrode techniques, the images are less direct observations of neural spike trains, and have limited time resolution and signal quality. To infer the functional connectivity matrix of observable micro-circuits from such indirect data, we assume a coupled hidden-Markov chain model, where each neuron is a generalized linear model. The sufficient statistics of the cross-coupling terms, which comprise the functional connectivity matrix, are obtained using an embedded-chain-within-blockwise-Gibbs algorithm for jointly sampling spike trains, given the calcium traces. By utilizing a factorized approximation, we implement our algorithm in parallel on a high-performance cluster, without a significant degradation of our results. Furthermore, by imposing biophysically realistic constraints on our model, such as a sparse constraint on the functional connectivity matrix, we can reduce the amount of data required to obtain a good fit. In realistic simulations, we show that our method can successfully infer connectivity patterns from ~ 10 minutes of calcium imaging data for neural populations up to 500 neurons large, in only ~ 10 minutes of computation time. Finally, we show the robustness of our method to various model misspecifications. Thus, this approach seems ready to be utilized by experimental neuroscientists to unravel the functional connection matrices underlying behavior and perception.

1 Introduction

Since Ramon y Cajal discovered that the brain is a rich and dense *network* of neurons (Ramon y Cajal, 1904; Ramon y Cajal, 1923), neuroscientists have been intensely curious about the details of these networks, which are believed to be the biological substrate for memory, cognition, and perception. While we have learned a great deal in the last century about "macro-circuits" — the connectivity between coarsely-defined brain areas — relatively little is known about "micro-circuits," i.e., the connectivity within populations of neurons at a fine-grained cellular level. Broadly, one can imagine two complementary strategies for inferring microcircuit connectivity: anatomical and functional. Anatomical approaches to inferring circuitry include any strategy that does not consider physiological measurements of neural activity: for example, recently developed technologies including array tomography (Micheva and Smith, 2007), "brainbow" mice (Livet et al., 2007), and serial electron microscopy (Briggman and Denk, 2006) are rapidly improving and show great promise. Our work, on the other hand, takes a functional approach: our aim is to infer microcircuit properties by observing the activity of a population of neurons.

Experimental tools that enable approximately simultaneous observations of the activity of many (e.g., $O(10^3)$) neurons are now widely available. While arrays of extracellular electrodes have been exploited for this purpose, the arrays most often used *in vivo* are inadequate for inferring monosynaptic connectivity, as the inter-electrode spacing is typically too large to record from all of the neurons in a given volume (Hatsopoulos et al., 1998; Harris et al., 2003; Stein et al., 2004; Santhanam et al., 2006; Luczak et al., 2007) [1]. Alternately, calcium-sensitive fluorescent indicators allow us to observe the spiking activity of many neighboring neurons (Tsien, 1989), which are more likely to be connected (Abeles, 1991; Braitenberg and Schuz, 1998). Some organic dyes achieve sufficiently high signal-to-noise ratios (SNR) that individual action potentials (spikes) may be resolved (Yuste et al., 2006), and bulk-loading techniques enable experimentalists to simultaneously fill populations of neurons with such

[1] It is worth noting, however, that multielectrode arrays which have been recently developed for use in the retina are capable of much denser sampling (Segev et al., 2004; Litke et al., 2004; Petrusca et al., 2007; Pillow et al., 2008).

dyes (Stosiek et al., 2003). In addition, genetically encoded calcium indicators are under rapid development in a number of groups, and are approaching SNR levels of nearly single spike accuracy as well (Wallace et al., 2008). Microscopy technologies for collecting fluorescence signals are also rapidly developing. Cooled CCDs for wide-field imaging (either epifluorscence or confocal) now achieve a quantum efficiency of $\approx 90\%$ with frame rates up to 60 Hz or greater, depending on the width of the field of view (Djurisic et al., 2004). For in vivo work, 2-photon laser scanning microscopy can achieve similar frame rates, using acoustic-optical deflectors to focus light at arbitrary locations in three-dimensional space (Reddy and Saggau, 2005; Iyer et al., 2006; Salome et al., 2006; Reddy et al., 2008), or using resonant scanners (Nguyen et al., 2001). Together, these experimental tools can provide movies of calcium fluorescence transients for small populations of neurons (e.g., $O(10^2)$), with "reasonable" SNR, at imaging frequencies of 30 Hz or greater, in both the in vitro and in vivo settings.

Given these experimental advances in functional neural imaging, our goal is to develop efficient computational and statistical methods to exploit this data for the analysis of neural connectivity. See Fig. 2 for a schematic overview. One major challenge is that calcium transients due to action potentials decay about an order of magnitude more slowly than the time course of the underlying neural activity (Yuste et al., 2006). Thus, to properly analyze the functional network connectivity in this setting, we must incorporate methods for effectively deconvolving the observed noisy fluorescence signal to obtain estimates of the underlying spiking rates (Yaksi and Friedrich, 2006; Greenberg et al., 2008; Vogelstein et al., 2009). To this end we introduce a coupled Markovian state-space model that relates the observed variables (fluorescence traces from the neurons in the microscope's field of view) to the hidden variables (spike trains and intracellular calcium concentrations of these neurons), as governed by a set of biophysical parameters including the network connectivity matrix. Given this model, we derive a Monte Carlo Expectation Maximization (MCEM) algorithm for approximating the maximum a posteriori estimates of the parameters of interest. Standard Monte Carlo sampling procedures (e.g., Gibbs sampling or sequential Monte Carlo) are inadequate in this setting, due to the high dimensionality and non-linear, non-Gaussian dynamics of the hidden variables in our model; we therefore develop a specialized blockwise-Gibbs approach to overcome these obstacles. This strategy enables us to accurately infer the functional connection matrix from moderately-sized simulated neural populations, under realistic assumptions about the dynamics and observation parameters. We describe our approach below, along with several methods for improving its computational speed and statistical efficiency.

2 Methods

2.1 Model

We begin by detailing a parametric generative model for the (unobserved) joint spike trains of all N observable neurons, along with the observed calcium fluorescence data. Each neuron is modeled as a generalized linear model (GLM); this class of model is known to capture the statistical firing properties of individual neurons fairly accurately (Brillinger, 1988; Chornoboy et al., 1988; Brillinger, 1992; Plesser and Gerstner, 2000; Pillow et al., 2008; Paninski et al., 2004; Paninski, 2004; Rigat et al., 2006; Truccolo et al., 2005; Nykamp, 2007; Kulkarni and Paninski, 2007; Vidne et al., 2009; Stevenson et al., 2009). We denote the i-th neuron's activity at time t as $n_i(t)$: in continuous time, $n_i(t)$ could be modeled as an unmarked point process, but we will take a discrete-time approach here, with each $n_i(t)$ taken to be a binary

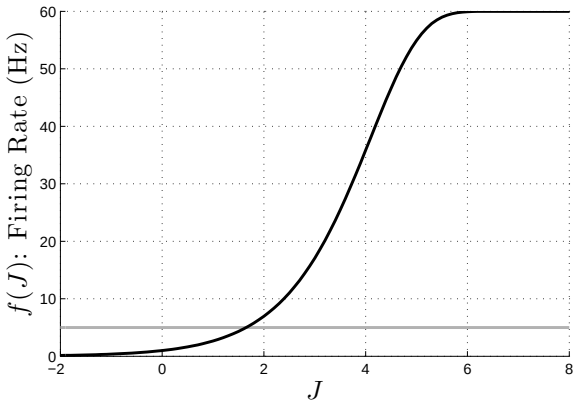

Figure 1: A plot of the firing rate nonlinearity $f(J)$ used in our simulations; see Eq. 1. Note that the firing rate saturates at $1/\Delta$, because of our Bernoulli assumption (i.e., the spike count per bin is at most one); $\Delta = (60 \text{ Hz})^{-1}$ here.

random variable. We model the spiking probability of neuron i via an instantaneous nonlinear function, $f(\cdot)$, of the filtered and summed input to that neuron at that time, $J_i(t)$. This input is composed of: (i) some baseline value, b_i; (ii) some external vector stimulus, $S(t)$, that is linearly filtered by k_i; and (iii) spike history terms, $h_{ij}(t)$, from each neuron j, weighted by w_{ij}:

$$n_i(t) \sim \text{Bernoulli}\big(f\big(J_i(t)\big)\big), \qquad J_i(t) = b_i + k_i \cdot S(t) + \sum_{j=1}^{N} w_{ij} h_{ij}(t). \tag{1}$$

To ensure computational tractability, we must impose some reasonable constraints on the instantaneous nonlinearity $f(\cdot)$ (which plays the role of the inverse of the link function in the standard GLM setting) and on the dynamics of the spike-history effects $h_{ij}(t)$. First, we restrict our attention to functions $f(\cdot)$ which ensure the concavity of the spiking loglikelihood in this model (Paninski, 2004; Escola and Paninski, 2008), as we will discuss at more length below. In this paper, we use

$$f(J) = P[n > 0 \mid n \sim Poiss(e^J \Delta)] = 1 - \exp[-e^J \Delta] \tag{2}$$

(Fig. 1), where the inclusion of Δ, the time step size, ensures that the firing rate scales properly with respect to the time discretization; see (Escola and Paninski, 2008) for a proof that this $f(\cdot)$ satisfies the required concavity constraints. However, we should note that in our experience the results depend only weakly on the details of $f(\cdot)$ within the class of log-concave models (Li and Duan, 1989; Paninski, 2004).

Second, because the algorithms we develop below assume Markovian dynamics, we model the spike history terms in terms of an autoregressive process:

$$h_{ij}(t) = (1 - \Delta/\tau_h) h_{ij}(t - \Delta) + n_j(t) + \sigma_h \sqrt{\Delta} \epsilon_{ij}(t), \tag{3}$$

where τ_h is the decay time constant for spike history terms, σ_h is a standard deviation parameter, $\sqrt{\Delta}$ ensures that the statistics of this Markov process have a proper Ornstein-Uhlenbeck limit as $\Delta \to 0$, and throughout this paper, ϵ denotes an independent standard normal random variable. Note that this model generalizes (via a simple augmentation of the state variable $h_{ij}(t)$) to allow each neuron to have several spike history terms, each with a

unique time constant, which when weighted and summed allow us to model a wide variety of possible post-synaptic effects, including bursting, facilitating, and depressing synapses; see (Vogelstein et al., 2009) for further details. We restrict our attention to the case of a single time constant τ_h here, both for simplicity and also because we find that the detailed shape of the coupling terms $h_{ij}(t)$ had a limited effect on the inference of the connectivity matrix, as illustrated in Fig. 12 below. In addition, we treat τ_h and σ_h as known synaptic parameters; therefore our unknown model spiking parameters are $\{\mathbf{w}_i, k_i, b_i\}$, with $\mathbf{w}_i = (w_{i1}, \ldots, w_{iN})$.

The problem of estimating the connectivity parameters $\mathbf{w}_i$ in this type of GLM, given a fully-observed ensemble of neural spike train $\{n_i(t)\}$, has recently received a great deal of attention; see the references above for a partial list. In the calcium fluorescent imaging setting, however, we do not directly observe spike trains; $\{n_i(t)\}$ must be considered a hidden variable here. Instead, each spike in a given neuron leads to a rapid increase in the intracellular calcium concentration, which then decays slowly due to various cellular buffering and extrusion mechanisms. We in turn make only noisy, indirect, and subsampled observations of this intracellular calcium concentration, via fluorescent imaging techniques (Yuste et al., 2006). To perform statistical inference in this setting, (Vogelstein et al., 2009) proposed a simple conditional first-order hidden Markov model (HMM) for the intracellular calcium concentration $C_i(t)$ in cell i at time t, along with the observed fluorescence $F_i(t)$:

$$C_i(t) = C_i(t-\Delta) + (C_i^b - C_i(t-\Delta))\Delta/\tau_i^c + A_i n_i(t) + \sigma_i^c \sqrt{\Delta}\epsilon_i^c(t), \tag{4}$$

$$F_i(t) = \alpha_i S(C_i(t)) + \beta_i + \sqrt{\gamma_i S(C_i(t)) + \sigma_i^F}\epsilon_i^F(t). \tag{5}$$

This model can be interpreted as a simple driven autoregressive process: under nonspiking conditions, $C_i(t)$ fluctuates around the baseline level of C_i^b, driven by normally-distributed noise $\epsilon_i^c(t)$ with standard deviation $\sigma_i^c\sqrt{\Delta}$. Whenever the neuron fires a spike, $n_i(t) = 1$, causing the calcium variable $C_i(t)$ to jump by a fixed amount A_i, and subsequently decay with time constant τ_i^c. The fluorescence signal $F_i(t)$ corresponds to the count of photons collected at the detector per neuron per imaging frame. This photon count may be modeled with normal statistics, with the mean and variance given by generalized Hill functions, $S(\cdot)$, where $S(C) = C/(C + K_d)$ (Yasuda et al., 2004). Because the parameter K_d effectively acts as a simple scale factor, and is a property of the fluorescent indicator, we assume throughout this work that it is known.

To summarize, Eqs. 1 – 5 define a coupled HMM: the underlying spike trains $n_i(t)$ and spike history terms $h_{ij}(t)$ evolve in a Markovian manner given the stimulus $S(t)$. These spike trains in turn drive the intracellular calcium concentrations $C_i(t)$, which are themselves Markovian, but evolving at a slower timescale τ_i^c. Finally, we observe only the fluorescence signals $\{F_i(t)\}$, which are related in a simple Markovian fashion to the calcium variables $C_i(t)$.

2.2 Goal and general strategy

Our primary goal is to estimate the connectivity matrix, $\mathbf{w}$, given the observed set of calcium fluorescence signals $\mathbf{F}$. We must also deal with a number of nuisance parameters: the spiking parameters $\{k_i, b_i\}$ and the calcium parameters $\{C_i^b, \tau_i^c, A_i, \sigma_i^c, \alpha_i, \beta_i, \gamma_i, \sigma_i^F\}$. We addressed the problem of estimating these latter parameters in earlier work (Vogelstein et al., 2009); thus our focus here will be on the connectivity matrix $\mathbf{w}$. A Bayesian approach is natural here, since we have a good deal of prior information about neural connectivity; see (Rigat et al., 2006) for a related discussion. However, a fully-Bayesian approach, in which we numerically integrate

over the very high-dimensional parameter $\theta = \{\mathbf{w}, k_i, b_i, C_i^b, \tau_i^c, A_i, \sigma_i^c, \alpha_i, \beta_i, \gamma_i, \sigma_i^F\}$, is not particularly attractive here, from a computational point of view. Thus we take a compromise approach and compute *maximum a posteriori* (MAP) estimates for the parameters via an expectation-maximization (EM) algorithm in which the sufficient statistics are computed by a hybrid blockwise Gibbs sampler and sequential Monte Carlo (SMC) method.

More specifically, we iterate the steps:

$$\textbf{E step: } \text{Evaluate } Q(\theta, \theta^{(l)}) = E_{P[\mathbf{X}|\mathbf{F};\theta^{(l)}]} \ln P[\mathbf{F}, \mathbf{X}|\theta] = \int P[\mathbf{X}|\mathbf{F};\theta^{(l)}] \ln P[\mathbf{F}, \mathbf{X}|\theta] d\mathbf{X}$$

$$\textbf{M step: } \text{Solve } \theta^{(l+1)} = \operatorname*{argmax}_{\theta} \left\{ Q(\theta, \theta^{(l)}) + \ln P(\theta) \right\},$$

where $\mathbf{X}$ denotes the set of all hidden variables $\{C_i(t), n_i(t), h_{ij}(t)\}_{i,j \le N, t \le T}$ and $P(\theta)$ denotes a (possibly improper) prior on the parameter space θ. According to standard EM theory (Dempster et al., 1977; McLachlan and Krishnan, 1996), each iteration of these two steps is guaranteed to increase the log-posterior $\ln P(\theta^{(l)}|\mathbf{F})$, and will therefore lead to at least a locally maximum a posteriori estimator.

Now our major challenge is to evaluate the auxiliary function $Q(\theta, \theta^{(l)})$ in the E-step. Because our model is a coupled HMM, as discussed in the previous section, Q simplifies considerably (Rabiner, 1989):

$$\begin{aligned} Q(\theta, \theta^{(l)}) &= \sum_{it} P[C_i(t)|\mathbf{F};\theta] \times \ln P[F_i(t)|C_i(t); \alpha_i, \beta_i, \gamma_i, \sigma_i^F] \\ &+ \sum_{it} P[C_i(t), C_i(t-\Delta), n_i(t)|\mathbf{F};\theta] \times \ln P[C_i(t)|C_i(t-\Delta), n_i(t); C_i^b, \tau_i^c, A_i, \sigma_i^c] \\ &+ \sum_{it} P[n_i(t), \mathbf{h}_i(t)|\mathbf{F};\theta] \times \ln P[n_i(t)|\mathbf{h}_i(t); b_i, k_i, \mathbf{w}_i, S(t)], \end{aligned} \tag{6}$$

where $\mathbf{h}_i(t) = \{h_{ij}(t)\}_{j \le N}$. Thus we need only compute low-dimensional marginals of the full posterior distribution $P[\mathbf{X}(t)|\mathbf{F};\theta]$; specifically, we need pairwise marginals, of the form $P[X_i(t), X_i(t-\Delta)|\mathbf{F};\theta]$. The high dimensionality of the hidden variable $\mathbf{X}$ necessitates the development of specialized blockwise Gibbs-SMC sampling methods, as we describe in sections 2.3 and 2.4 below. Once we have obtained these marginals, the M-step breaks up into a number of independent optimizations that may be computed in parallel and which are therefore relatively straightforward (section 2.5); see section 2.6 for a pseudocode summary along with some specific implementation details.

2.3 Initialization of "intrinsic" parameters via sequential Monte Carlo methods

We begin by constructing relatively cheap, approximate preliminary estimators for the nuisance parameters $\theta \backslash \{w_{ij}\}_{i \neq j}$, i.e., the observation parameters, $\{\alpha_i, \beta_i, \gamma_i, \sigma_i^F\}$, the calcium dynamics parameters $\{C_i^b, \tau_i^c, A_i, \sigma_i^c\}_i$, and the intrinsic spiking parameters, $\{k_i, b_i, w_{ii}\}$. The idea is to initialize our estimator by assuming that each neuron is observed independently. Thus we want to compute $P[X_i(t), X_i(t-\Delta)|\mathbf{F}_i;\theta_i]$ (where $\mathbf{F}_i$ denotes the i-th observed fluorescence signal), and solve the M-step for each individual parameter θ_i (where θ_i includes all of neuron i's intrinsic parameters but excludes the connectivity parameters $\{w_{ij}\}_{i \neq j}$), with the connection matrix parameter held fixed. This single-neuron case is much simpler,

and has been discussed at length in (Vogelstein et al., 2009); therefore, we only provide a brief overview here. The standard forward and backward recursions provide these posteriors (Shumway and Stoffer, 2006):

$$P[X_i(t)|F_i(0:t)] \propto P[F_i(t)|X_i(t)] \int P[X_i(t)|X_i(t-\Delta)]P[X_i(t-\Delta)|F_i(0:t-\Delta)]dX_i(t-\Delta), \ for\ t = \Delta, 2\Delta, \ldots, T \tag{7}$$

$$P[X_i(t), X_i(t-\Delta)|\mathbf{F}_i] = P[X_i(t)|\mathbf{F}_i] \frac{P[X_i(t)|X_i(t-\Delta)]P[X_i(t-\Delta)|F_i(0:t-\Delta)]}{\int P[X_i(t)|X_i(t-\Delta)]P[X_i(t-\Delta)|F_i(0:t-\Delta)]dX_i(t-\Delta)}, \ for\ t = T, T-\Delta, \ldots, \Delta \tag{8}$$

where we have dropped the conditioning on the parameters θ for brevity's sake. Because these integrals cannot be analytically evaluated for our model, we approximate them using SMC ("marginal particle filtering") methods (Doucet et al., 2000; Doucet et al., 2001; Godsill et al., 2004); see (Vogelstein et al., 2009) for details on the proposal density and resampling methods used here. The output of these SMC techniques comprise an array of particle positions $\{X_i^{(m)}(t)\}$, where m indexes the particle number, and a discrete approximation to the marginals $P[X_i(t), X_i(t-\Delta)|\mathbf{F}_i, \theta_i]$,

$$P[X_i(t), X_i(t-\Delta)|\mathbf{F}_i, \theta_i] \approx \sum_{m,m'} r_i^{(m,m')}(t, t-\Delta)\delta\left[X_i(t) - X_i^{(m)}(t)\right] \times \delta\left[X_i(t-\Delta) - X_i^{(m')}(t-\Delta)\right], \tag{9}$$

where $r_i^{(m,m')}(t, t-\Delta)$ denotes the weight attached to the pair of particles with positions $\left(X_i^{(m)}(t), X_i^{(m')}(t-\Delta)\right)$, and $\delta(.)$ denotes a Dirac mass.

As discussed above, the sufficient statistics for estimating the intrinsic parameters for each neuron, θ_i, are exactly these marginal posteriors. As shown in Eq. 6, the M-step decouples into three independent subproblems. The first term depends on only $\{\alpha_i, \beta_i, \gamma_i, \sigma_i\}$; since $\ln P[F_i(t)|C_i(t); \theta_i]$ is quadratic (by our Gaussian assumption on the fluorescent observation noise), we can estimate these parameters by solving a weighted regression problem (specifically, we use a coordinate-optimization approach: we solve a quadratic problem for $\{\alpha_i, \beta_i\}$ while holding $\{\gamma_i, \sigma_i\}$ fixed, then estimate $\{\gamma_i, \sigma_i\}$ by the usual residual error formulas while holding $\{\alpha_i, \beta_i\}$ fixed). Similarly, the second term requires us to optimize over $\{\tau_i^c, A_i, C_i^b\}$, and then we use the residuals to estimate σ_i^c. Note that all the parameters mentioned so far are constrained to be non-negative, but may be solved efficiently using standard quadratic program solvers if we use the simple reparameterization $\tau_i^c \to 1 - \Delta/\tau_i^c$. Finally, the last term, assuming neurons are independent, may be expanded:

$$E[\ln P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta]] = P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}] \ln f(J_i(t)] + (1 - P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}]) \ln[1 - f(J_i(t))]; \tag{10}$$

since $J_i(t)$ is a linear function of $(b_i, k_i, \mathbf{w}_i)$, and the right-hand side of (10) is concave in $J_i(t)$, we see that the third term in Eq. (6) is a sum of terms which are concave in $(b_i, k_i, \mathbf{w}_i)$ — and therefore also concave in the linear subspace (b_i, k_i, w_{ii}) with $\{w_{ij}\}_{i \neq j}$ held fixed — and may thus be maximized efficiently using any convex optimization method, e.g. Newton-Raphson or conjugate gradient ascent.

Our procedure therefore is to initialize the parameters for each neuron using some default values that we have found to be practically effective in analyzing real data, and then recursively (i) estimate the marginal posteriors via Eq. (9) (E step), and (ii) maximize the intrinsic

parameters θ_i (M step), using the above described approach. We iterate these two steps until the change in parameters does not exceed some minimum threshold. We can then use the marginal posteriors from the last iteration to seed the blockwise Gibbs sampling procedure described below, to obtain a rough estimate of $P[\mathbf{h}_i(t), n_i(t)|\mathbf{F}]$.

2.4 Estimating joint posteriors over weakly coupled neurons

Now we turn to the key problem: computing $P(\mathbf{h}_i(t), n_i(t)|\mathbf{F}, \theta)$, which encapsulates the sufficient statistics for estimating the connectivity matrix $\mathbf{w}$ (recall equation (6)). The SMC methods described in the preceding section only provide the marginals over each neuron, $P[X_i(t), X_i(t-\Delta)|\mathbf{F}_i; \theta_i]$; these methods may in principle be extended to obtain the desired full posterior $P[\mathbf{X}(t), \mathbf{X}(t-\Delta)|\mathbf{F}; \theta]$, but the SMC algorithm is fundamentally a sequential importance sampling method, and therefore scales poorly as the dimensionality of the hidden state $\mathbf{X}(t)$ increases (Bickel et al., 2008). Thus we need a different approach.

One very simple idea is to use a Gibbs sampler: sample sequentially from

$$X_i(t) \sim P[X_i(t)|\mathbf{X}_{\setminus i}, X_i(0), \ldots, X_i(t-\Delta), X_i(t+\Delta), \ldots, X_i(T), \mathbf{F}; \theta], \tag{11}$$

looping over all cells i and all time bins t. Unfortunately, this approach is likely to mix poorly, due to the strong temporal dependence between $X_i(t)$ and $X_i(t+\Delta)$. Instead, we propose a blockwise Gibbs strategy, sampling each spike train as a block:

$$X_i \sim P[X_i|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta]; \tag{12}$$

if we can draw these blockwise samples $X_i = \{X_i(t)\}$ efficiently for a large subset of adjacent timebins t simultaneously, then we would expect the resulting Markov chain to mix much more quickly than the single-element Gibbs chain, since by assumption the hidden variables $\mathbf{X}_i, \mathbf{X}_j$ are weakly dependent for different cells $i \neq j$, and Gibbs is most efficient for weakly-dependent variables (Robert and Casella, 2005).

So, how can we efficiently sample from $P[\mathbf{X}_i|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta]$? One attractive approach is to try to repurpose the SMC methods described above, which are quite effective for drawing approximate samples from $P[\mathbf{X}_i|\mathbf{X}_{\setminus i}, \mathbf{F}_i; \theta]$ for one neuron i at a time. Recall that sampling from an HMM is in principle easy by the "propagate forward, sample backward" method: we first compute the forward probabilities $P[X_i(t)|\mathbf{X}_{\setminus i}(0:t), F_i(0:t); \theta]$ recursively for timesteps 0 up to T, then sample backwards from $P[X_i(t)|\mathbf{X}_{\setminus i}(0:T), F_i(0:T), X_i(t+\Delta); \theta]$. This approach is powerful because each sample requires just linear time to compute (i.e., $O(T/\Delta)$ time, where T/Δ is the number of desired time steps). Unfortunately, in this case we can only compute the forward probabilities approximately (with the SMC forward recursion (7)), and so therefore this attractive forward-backward approach only provides approximate samples from $P[\mathbf{X}_i|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta]$, not the exact samples required to establish the validity of the Gibbs method.

Of course, in principle we should be able to use the Metropolis-Hastings (M-H) algorithm to correct these approximate samples. The problem is that the M-H acceptance ratio in this setting involves a high-dimensional integral over the set of paths that the particle filter might possibly trace out, and is therefore difficult to compute directly. (Andrieu et al., 2007) discuss this problem at more length, along with some proposed solutions. However, a slightly simpler approach was introduced by (Neal et al., 2003). Their idea is to exploit the $O(T/\Delta)$ forward-backward sampling method by embedding a discrete Markov chain within the continuous state

space $\mathcal{X}_t$; the state space of this discrete embedded chain is sampled randomly according to some distribution ρ_t with support on $\mathcal{X}_t$. It turns out that an appropriate acceptance probability (defined in terms of the original state space model transition and observation probabilities, along with the auxiliary sampling distributions ρ_t) may be computed quite tractably, guaranteeing that the samples produced by this algorithm form a Markov chain with the desired equilibrium density. See (Neal et al., 2003) for details.

We can apply this embedded-chain method quite directly here to sample from $P[\mathbf{X}_i|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta]$. The one remaining question is how to choose the auxiliary densities ρ_t. We would like to choose these densities to be close to the desired marginal densities $P[X_i(t)|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta]$, and conveniently, we have already computed a good (discrete) approximation to these densities, using the SMC methods described in the last section. The algorithm described in (Neal et al., 2003) requires that ρ_t be continuous densities, so we simply convolve our discrete SMC-based approximation (specifically, the $X_i(t)$-marginal of Eq. (9)) with an appropriate normal density to arrive at a very tractable mixture-of-Gaussians representation for ρ_t.

Thus, to summarize, our procedure for sampling from the desired joint state distributions $P(\mathbf{h}_i(t), n_i(t)|F, \theta)$ has a Metropolis-within-blockwise-Gibbs flavor, where the internal Metropolis step is replaced by the $O(T/\Delta)$ embedded-chain method introduced by (Neal et al., 2003), and the auxiliary densities ρ_t necessary for implementing the embedded-chain sampler are obtained using the SMC methods from (Vogelstein et al., 2009).

2.4.1 A factorized approximation of the joint posteriors

If the SNR in the calcium imaging is sufficiently high, then by definition the observed fluorescence data F_i will provide enough information to determine the underlying hidden variables X_i. Thus, in this case the joint posterior approximately factorizes into a product of marginals for each neuron i:

$$P[\mathbf{X}|\mathbf{F}; \theta] \approx \prod_{i=1}^{N} P[\mathbf{X}_i|\mathbf{F}_i; \theta]. \tag{13}$$

We can take advantage of this representation because we have already estimated all the above marginals using the SMC methods described in section 2.3. In particular, we can obtain the sufficient statistics $P(\mathbf{h}_i(t), n_i(t)|\mathbf{F}, \theta)$ by forming a product over the marginals $P(X_i(t)|\mathbf{F}_i, \theta)$ obtained from (9). This factorized approximation entails a very significant gain in efficiency for two reasons: first, it obviates the need to generate joint samples via the expensive blockwise-Gibbs approach described above; and second, because we can very easily parallelize the SMC step, inferring the marginals $P[X_i(t)|\mathbf{F}_i; \theta]$ and estimating the parameters $(\theta_i, \mathbf{w}_i)$ for each neuron on a separate processor. We will discuss the empirical accuracy of this approximation in more depth in the Results section.

2.5 Estimating the functional connectivity matrix

Computing the M-step for the connectivity matrix, $\mathbf{w}$, is an optimization problem with on the order of N^2 variables. The auxiliary function (6) is concave in $\mathbf{w}$, and decomposes into N separable terms that may be optimized independently using standard ascent methods. To improve our estimates, we will incorporate two sources of strong *a priori* information via our prior $P(\mathbf{w})$: first, previous anatomical studies have established that connectivity in many neuroanatomical substrates is "sparse," i.e., most neurons form synapses with only a fraction

of their neighbors (Buhl et al., 1994; Thompson et al., 1988; Reyes et al., 1998; Feldmeyer et al., 1999; Gupta et al., 2000; Feldmeyer and Sakmann, 2000; Petersen and Sakmann, 2000; Binzegger et al., 2004; Song et al., 2005; Mishchenko et al., 2009), implying that many elements of the connectivity matrix $\mathbf{w}$ are zero; see also (Paninski, 2004; Rigat et al., 2006; Pillow et al., 2008; Stevenson et al., 2008) for further discussion. Second, "Dale's law" states that each of a neuron's postsynaptic connections in adult cortex (and many other brain areas) must all be of the same sign (either excitatory or inhibitory). Both of these priors are easy to incorporate in the M-step optimization, as we discuss below.

2.5.1 Imposing a sparse prior on the functional connectivity

It is well-known that imposing sparseness via an $L1$-regularizer can dramatically reduce the amount of data necessary to accurately reconstruct high-dimensional parameters (Tibshirani, 1996; Tipping, 2001; Donoho and Elad, 2003; Ng, 2004; Candes and Romberg, 2005; Mishchenko, 2009). We incorporate a prior of the form $\ln p(\mathbf{w}) = const. - \lambda \sum_{i,j} |w_{ij}|$, and additionally enforce the constraints $|w_{ij}| < m$, for a suitable constant m (since both excitatory and inhibitory cortical connections are known to be bounded in size). Since the penalty $\ln p(\mathbf{w})$ is concave, and the constraints $|w_{ij}| < m$ are convex, we may still solve the resulting optimization problem in the M-step using standard convex optimization methods (Boyd and Vandenberghe, 2004). In addition, the problem retains its separable structure: the full optimization may be broken up into N smaller problems that may be solved independently.

2.5.2 Imposing Dale's law on the functional connectivity

Enforcing Dale's law requires us to solve a non-convex, non-separable problem: we need to optimize the concave function $Q(\theta, \theta^{(l)}) + \ln P(\theta)$ under the non-convex, non-separable constraint that all of the elements in any column of the matrix $\mathbf{w}$ are of a definite sign (either nonpositive or nonnegative). It is difficult to solve this nonconvex problem exactly, but we have found that simple greedy methods are quite efficient in finding good approximate solutions. We begin with our original sparse solution, obtained as discussed in the previous subsection without enforcing Dale's law. Then we assign each neuron as either excitatory or inhibitory, based on the weights we have inferred in the previous step: i.e., neurons i whose inferred postsynaptic connections w_{ij} are largely positive are tentatively labeled excitatory, and neurons with largely inhibitory inferred postsynapic connections are labeled inhibitory. Neurons which are highly ambiguous may be unassigned in the early iterations, to avoid making mistakes from which it might be difficult to recover. Given the assignments a_i ($a_i = 1$ for putative excitatory cells, -1 for inhibitory, and 0 for neurons which have not yet been assigned) we solve the convex, separable problem

$$\underset{a_i w_{ij} \geq 0, |w_{ij}| < m \ \forall i,j}{\operatorname{argmax}} \ Q(\theta, \theta^{(l)}) - \lambda \sum_{ij} |w_{ij}| \tag{14}$$

which may be handled using the standard convex methods discussed above. Given the new estimated connectivities $\mathbf{w}$, we can re-assign the labels a_i, or even flip some randomly to check for local optima. We have found this simple approach to be effective in practice.

Algorithm 1 Pseudocode for estimating functional connectivity from calcium imaging data using EM; η^n, η^F, N_G are user-defined convergence tolerance parameters.

while $|\mathbf{w}^{(l)} - \mathbf{w}^{(l-1)}| > \eta^w$ **do**
  for all $i = 1 \ldots N$ **do**
    while $|{\theta_i}^{(l)} - {\theta_i}^{(l-1)}| > \eta^F$ **do**
      Approximate $P[X_i(t)|\mathbf{F}_i; \theta]$ using SMC (section 2.3)
      Perform the M-step for the "intrinsic" parameters $\theta \backslash \{w_{ij}\}_{i \neq j}$ (section 2.3)
    end while
  end for
  for all $i = 1 \ldots N$ **do**
    Approximate $P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta]$ using either the blockwise Gibbs method or the high-SNR factorized approximation (section 2.4)
  end for
  for all $i = 1 \ldots N$ **do**
    Perform the M-step using separable convex optimization methods (section 2.5)
  end for
end while

2.6 Specific implementation notes

Pseudocode summarizing our approach is given in Algorithm 1. As discussed in section 2.3, the "intrinsic" parameters $\theta \backslash \{w_{ij}\}_{i \neq j}$ may be initialized effectively using the methods described in (Vogelstein et al., 2009); then the full parameter θ is estimated via EM, where we use the embedded-chain-within-blockwise-Gibbs approach discussed in section 2.4 (or the cheaper factorized approximation described in section 2.4.1) to obtain the sufficient statistics in the E step and the separable convex optimization methods discussed in section 2.5 for the M step.

As emphasized above, the parallel nature of these EM steps is essential for making these computations tractable. We performed the bulk of our analysis on a 100-node cluster of Intel Xeon L5430 based computers (2.66 GHz). For 10 minutes of simulated fluorescence data, imaged at 30 Hz, calculations using the factorized approximation typically took 10-20 minutes per neuron (divided by the number of available processing nodes on the cluster), with time split approximately equally between (i) estimating the intrinsic parameters $\theta \backslash \mathbf{w}_{ij}$, (ii) approximating the posteriors using the independent SMC method, and (iii) estimating the functional connectivity matrix, $\mathbf{w}$. The hybrid embedded-chain-within-blockwise-Gibbs sampler was substantially slower, up to an hour per neuron per Gibbs pass, with the Gibbs sampler dominating the computation time, because we thinned the chain by a factor of five (since we found empirically that the autocorrelation of the Gibbs chain had a scale of about five time steps).

3 Results

3.1 Simulating neural activity in a neural population

To test the described method for inferring functional connectivity from calcium imaging data, we simulated networks (according to our model, Eqs. 1 – 5) of stochastically connected

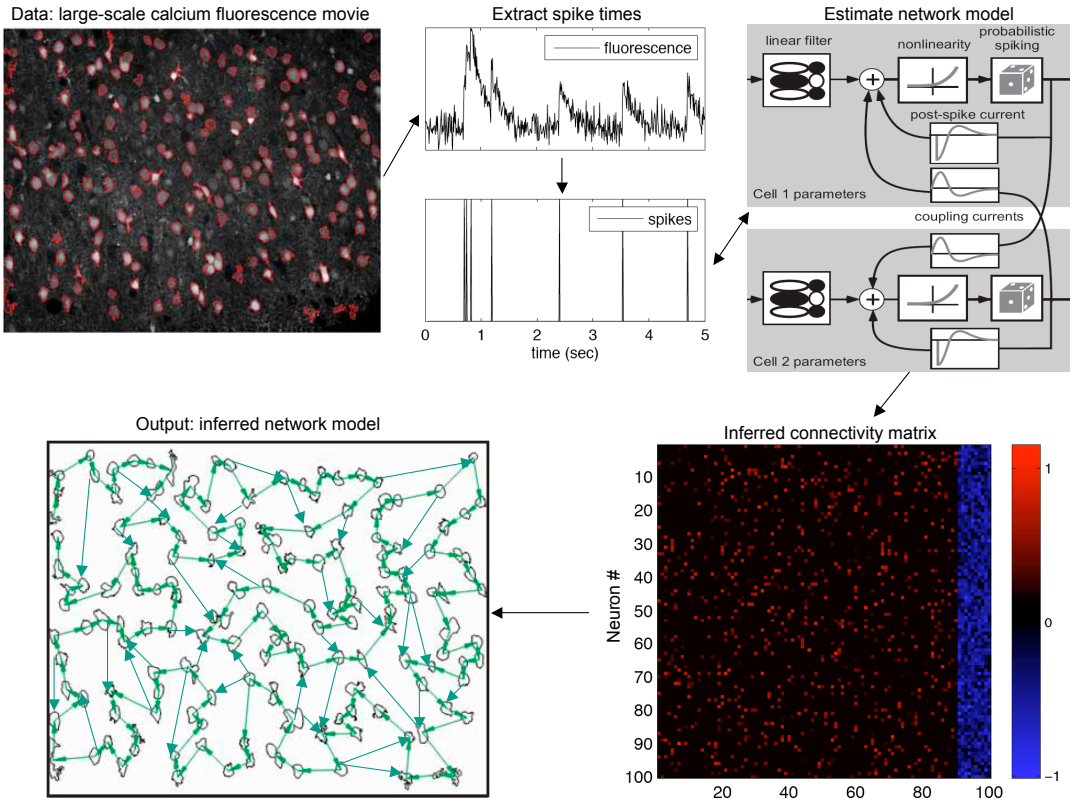

Figure 2: Schematic overview. The raw observed data is a large-scale calcium fluorescence movie, which is pre-processed to correct for movement artifacts (in the in vivo setting) and find regions-of-interest (i.e., putative neurons); note that we have omitted details of these important preprocessing steps in this paper. Given the fluorescence traces from each neuron, we estimate the underlying spike trains (i.e., time series of neural activity) using statistical deconvolution methods. Then we estimate the parameters of a network model, given the observed data. Our major goal is to obtain an accurate estimate of the network connectivity matrix, which summarizes the information we are able to infer about the local neuronal microcircuit. This figure adapted from personal communications with Rafael Yuste, Brendon O. Watson, and Adam Packer.

neurons. Although simulations ran at 1 msec time discretization, imaging rate was assumed to be much slower. Simulations lasted anywhere between 5 minutes and 1 hours (of simulated time). Model parameters were chosen based on experimental data available from the literature (Braitenberg and Schuz, 1998; Gomez-Urquijo et al., 2000; Lefort et al., 2009; Sayer et al., 1990).

More specifically, the network was divided into excitatory (80%) and inhibitory (20%) neurons (Braitenberg and Schuz, 1998; Gomez-Urquijo et al., 2000), each respecting Dale's law, i.e., each neurons was either excitatory or inhibitory (corresponding to all positive or all negative columns in our functional connection weight matrix, $\mathbf{w}$). Neurons were randomly

connected to each other with probability 0.1 (Braitenberg and Schuz, 1998; Lefort et al., 2009). Synaptic weights for excitatory connections, as defined by EPSP peak amplitude, were randomly drawn from exponential distribution with the mean of $0.5\mu V$ (Lefort et al., 2009; Sayer et al., 1990). To convert these into functional connectivity weights, we note that $w_{ij} = 1$ corresponds to a jump in $J_i = 1$, immediately following a spike from neuron j. Defining V_δ as the difference between threshold and rest for a neuron, and V_E as the peak amplitude of an EPSP (and similarly for IPSP's), we can say the change in probability after a spike is $\Delta P = V_E/V_\delta$. Because Eq. 1 provides the same quantity in our model parameter, w_{ij}, we have:

$$w_{ij} = \ln(-\ln(e^{-r_i\tau_w} - V_E/V_\delta)/r_i\tau_w), \tag{15}$$

where $r_i = \exp(b_i)$ is the base firing rate of neuron i and $\tau_w = 10$ msec was the typical EPSP/IPSP scale over which single EPSP affects the firing probability of the neuron i (). Inhibitory connections were also drawn from exponential distribution; their strengths chosen so as to balance excitatory and inhibitory currents in the network, and achieve an average firing rate of ≈ 5 Hz (**?**). Practically, the mean strength of inhibitory connections was about 10 times larger than that of the excitatory connections. PSP shapes were modeled as an alpha function (Koch, 1999), by differencing of two exponentials, corresponding to a sharp rise and relatively slow decay (Sayer et al., 1990). We neglected conduction delays, given that the time delay below ~ 1 msec expected in local cortical circuit was smaller than the time step of our computer simulation. Each neuron also had an exponential refractory current with a 10 msec time constant (Koch, 1999). XXX Y: did that not get sampled from some distribution? XXX

Parameters for the calcium dynamics and observation statistics were chosen according to our experience with several cells analyzed using algorithm of (Vogelstein et al., 2009), and conformed to other experimental observations (Yuste et al., 2006; Helmchen et al., 1996; Brenowitz and Regehr, 2007). Each parameter was generated from a normal distribution with specified mean and variance at about 30% of the mean, truncated at the lower bound at about 30% of the mean value. Table 1 provides details for each of the parameters in our model.

3.2 Inferring functional connectivity from the simulated calcium imaging data

With neural population activity prepared as described in the previous section, we used our inference algorithms to reconstruct the functional connectivity matrix from simulated fluorescence data. Specifically, we estimated the connectivity matrix by maximizing $E[\ln P[n_i(t)|\mathbf{h}(t); b_i, \mathbf{w}_i]]$ with respect to $\{b_i, \mathbf{w}_i\}$, for each neuron (c.f. Eq. 6), using both the embedded-chain-within-blockwise-Gibbs approach as well as using factorized approximation, Figure 4. We found that factorized approximation algorithm was able to provide reconstructions almost as accurate as the exact embedded-chain-within-blockwise-Gibbs approach — $r^2 = 0.47$ versus $r^2 = 0.48$ — when parameters corresponded to a realistic (but relatively high quality) preparation, given 10 minutes simulation time, in a population of $N = 25$ neurons. Figure 3 depicts a couple different fluorescence traces, of varying SNR. factorized

To compare the quality of our inferences using only the fluorescence traces, we also estimated the weights using the true spike trains, down-sampled to the frame rates of calcium

imaging. Indeed, the quality of our estimates using the fluorescence traces are worse than those obtained from the down-sampled spike trains; $r^2 = 0.57$ using the down-sampled spike trains. Note that $r^2 \to 1$ as $T \to \infty$, when using the true (i.e., not down-sampled) spike trains, as guaranteed by our model. Indeed, when using the true spike trains, $r^2 = XXX$ for this simulation. This suggests that as imaging rates increase, we can expect a corresponding increase in accuracy of our functional connectivity weight estimates.

3.3 Scale bias in inferred connection weights due to coarse time discretization of calcium imaging data

That all three approaches considered in Figure 4 exhibit a significant scaling bias suggests that down-sampling spike times introduces this bias. We conjectured that this is caused by the discrepancy between $\Delta \approx 15 - 30$ msec and the assumed PSP time scale $\approx 10 - 20$ msec. Properly estimating the magnitude of the connection weights w_{ij} is based on empirically evaluating the spike-triggered probability of neuron i to fire, conditioned on neuron j. Note that most triggered spikes of neuron i will occur within $\approx 10 - 20$ msec from a triggering spike of neuron j (because of our assumed values for τ_w). However, when spike trains are discretized into large time-bins, e.g. $\Delta = 30$ msec, a significant fraction of such triggered spike pairs seem coincident, as opposed to causal (because both the triggering and triggered spikes end up in the same time-bin). Therefore, we empirically observe a smaller number of triggered spikes than occur, causing a scale bias in w_{ij}.

To quantitatively estimate the magnitude of this scale bias and correct it, consider two neurons coupled with small weight w_{12}. Assume that, aside from this coupling, these neurons fire with baseline firing rate of $r = f(b) = \exp(b)$, $b \gg w_{12}$. In order to estimate w_{12} we observe the number of spike pairs such that the first neuron fired after the second neuron over some small period of time $\tau_w \ll t' \ll 1/r$, which is in excess of the baseline firing rate

Table 1: Table of simulation parameters. $\mathcal{E}(\lambda)$ indicates an exponential distribution with λ, and $\mathcal{N}(\mu, \sigma^2)$ indicates a truncated normal with mean μ and variance σ^2, and lower bound of one standard deviation below the mean.

Total neurons	10-500	
Excitatory neurons	80%	
Connections sparseness	10%	
Baseline firing rate	5	Hz
EPSP peak height	$\sim \mathcal{E}(0.5)$	μV
IPSP peak height	$\sim -\mathcal{E}(2.3)$	μV
EPSP rise time	$\sim \mathcal{N}(1, 0.3)$	msec
EPSP decay time	$\sim \mathcal{N}(10, 2.5)$	msec
IPSP rise time	$\sim \mathcal{N}(1, 0.3)$	msec
IPSP decay time	$\sim \mathcal{N}(20, 7)$	msec
Calcium std. σ_c	$\sim \mathcal{N}(28, 10)$	μM
Calcium jump after spike, A_c	$\sim \mathcal{N}(80, 20)$	μM
Calcium baseline, C_b	$\sim \mathcal{N}(24, 8)$	μM
Calcium decay time, τ_c	$\sim \mathcal{N}(500, 170)$	msec
Mean photon budget α_c	1–80	Kph/neuron/frame
Dissociation constant, K_d	200	μM

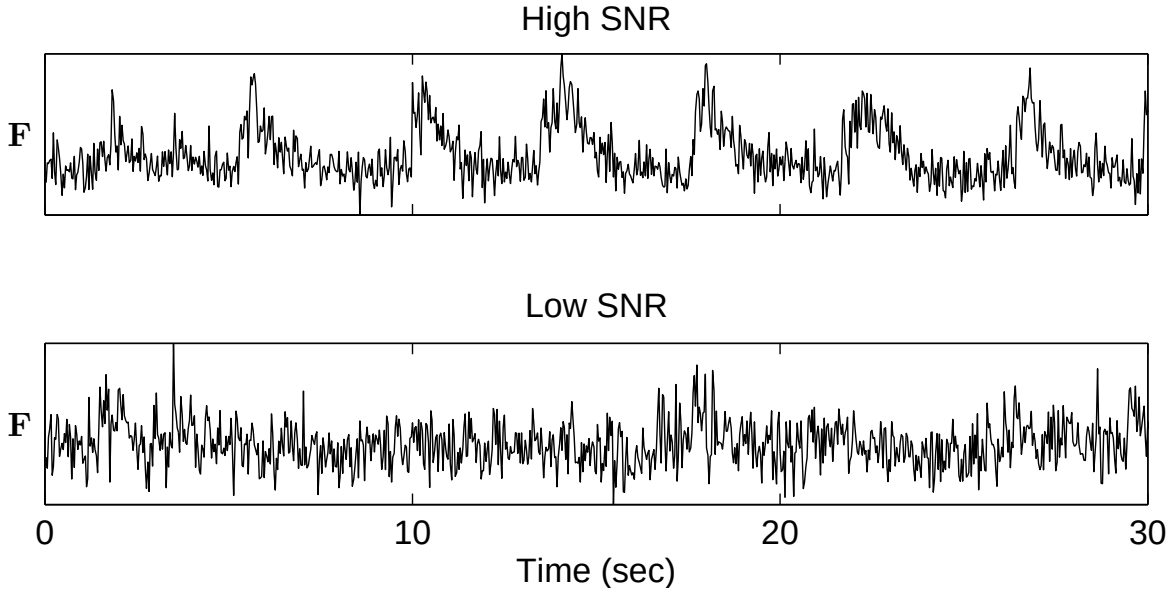

Figure 3: Two in vivo fluorescence traces (top) and their corresponding inferred spike trains (bottom). The left panels show a high SNR case, that is sufficient quality for the factorized approximation to work nearly as well as the embedded-chain-within-Gibbs approach. The right panels show a low SNR case, in which the factorized approach is insufficient. The ordinate is in arbitrary units. Data from the laboratory of Tom Mrsic-Flogel.

$\Delta n(2 \to 1) = n(2 \to 1) - rt' \approx f'(b) \int w_{12} \exp(-t/\tau_w)\ \mathrm{dt}$.

Now, assume that spike trains were additionally discretized into time-bins of size Δ, and only the spikes that occurred in different time-bins $t' < t$ were considered as non-coincidental. In this case, the number of spike pairs such that the spike of the first neuron followed that of the second neuron observed empirically is:

$$\Delta n'(2 \to 1) \approx f'(b) \int_0^{\Delta} \frac{1}{\Delta} \int_{\Delta}^{T} \exp(-(t_2 - t_1)/\tau_w) \mathrm{dt}_1 \mathrm{dt}_2. \tag{16}$$

The ratio of this empirical count, $\Delta n'(2 \to 1)$, to that expected from our model for given w_{12}, $\Delta n(2 \to 1)$, is the theoretical scale bias:

$$\Delta n'_{12} / \Delta n_{12} \approx \frac{1 - \exp(-\Delta/\tau_w)}{\Delta/\tau_w}. \tag{17}$$

In Figure 5 we plotted this estimated magnitude versus that empirically observed from our simulations for different values of Δ. As can be seen from Figure 5, simple theoretical estimate Eq. 17 describes observed scale bias quite well. This scale bias could potentially be overcome by inferring spike trains using the superresolution feature of (**?**), i.e., sampling spikes with a finer temporal resolution than the image rate.

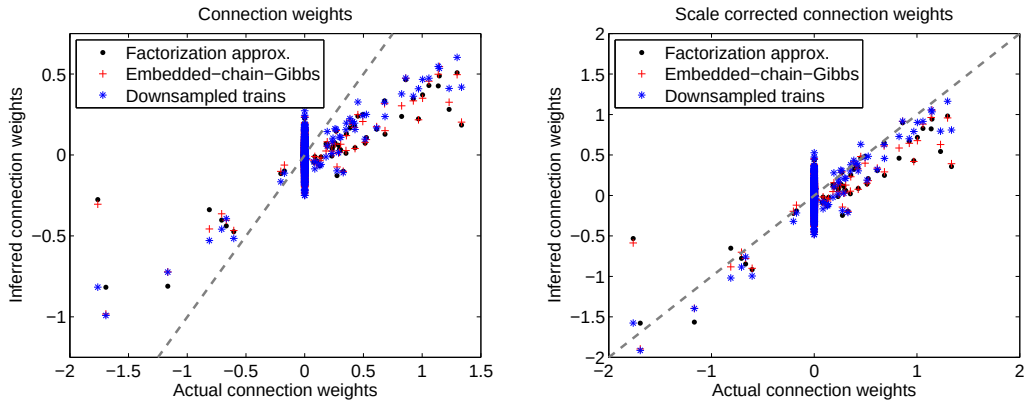

Figure 4: Functional connectivity matrix can be reconstructed from calcium imaging data. In the upper panels inferred connection weights are shown in a scatter plot versus real connection weights, with inference performed using factorized approximation algorithm, exact embedded-chain-within-blockwise-Gibbs approach, and original spike trains observed at the frame rate of the calcium imaging. Network of $N = 25$ neurons was used, firing at ≈ 5 Hz, and imaged for T=600 sec at intermediate SNR (photon budget 10Kph/neuron/frame, see below). $r^2 = 0.47$ for factorized approximation algorithm was found, $r^2 = 0.48$ for embedded-chain-within-blockwise-Gibbs approach, and $r^2 = 0.57$ for the original spike trains. Thus, factorized approximation produced results almost as accurate as the exact embedded-chain-within-blockwise-Gibbs approach, and almost as accurate as the original spikes. Inferred connectivity weights (upper left) were scaled with respect to true connectivity by a constant amount due to time discretization bias (see below); other than scale, inferred connectivity represented the true connectivity matrix very well (upper right). Thus, calcium imaging is sufficient to identify connected pairs of neurons reliably.

3.4 Impact of different imaging frame rates, noise levels, and durations on the estimator accuracy

What minimal conditions for the experimental setup should be met to allow for successful reconstruction of the connectivity from calcium imaging data? In Figures 6 – 8 we address this question. Figure 6 shows the quality of the inferred connectivity matrix as function of the imaging frame rate; imaging frame rates 30-60 Hz are needed to achieve meaningful reconstruction results. These imaging frame rates are feasible for already existing experimental setups (Nguyen et al., 2001; Reddy and Saggau, 2005; Iyer et al., 2006; Salome et al., 2006; Reddy et al., 2008). Figure 7 shows the quality of the inferred connectivity matrix as function of effective SNR and photon budget. Operationally, we define effective SNR here as $eSNR =$ XXX and photon budget as $1/\gamma$XXX. Note, therefore, that the relationship between $eSNR$ and $1/\gamma$ is XXX. From our experience with the analysis of real cells (Vogelstein et al., 2009), the eSNR in real data was $\sim XXX$ for in vivo data collected at 15 Hz and $\sim XXX$ for in vitro data at the same frame rate. As can be seen from Figure 7, the effective SNR necessary for accurate reconstructions was XXX. For lower photon counts, the amount of noise in calcium imaging data degraded inferred connectivity matrices significantly. Finally, Figure 8 shows the quality of the inferred connectivity matrix as function of the experimental duration. The minimal duration for a particular r^2 depended substantially on whether prior information about the distribution of connectivity weights was incorporated into the M-step

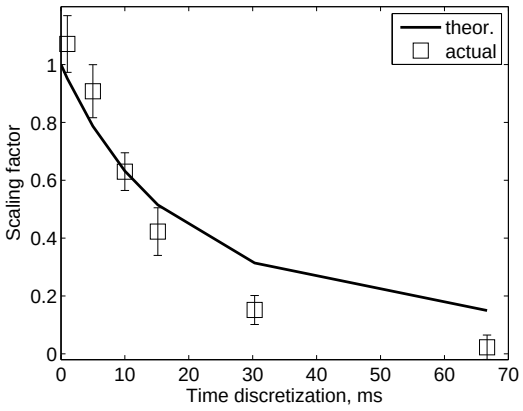

Figure 5: The low-frame rate of calcium imaging can explain the observed scale bias from Figure 4. Our theoretical scale bias is determined by considering what fraction of spikes would fall within a single image frame, which is a function of Δ (c.f. Eq. 17). The center of each square indicates the mean scale bias of 5 simulations, the extent of each square indicates one standard deviation, and the errorbars indicate the 95th and 5th percentiles. XXX Y: can you correct these details? XXX

(as described in detail below). In particular, for the M-step lacking a sparse prior, the calcium imaging duration necessary to achieve $r^2 = 0.5$ for the reconstructed connectivity matrix was $T \approx 10$ min, whereas, for the same data, when incorporating a sparse prior, $r^2 > 0.5$ was achieved already at $T \sim 5$ min. These numbers appear to be well within limitations of the existing experimental preparations. Furthermore, in agreement with theoretical analysis in Section 2, the accuracy of the reconstruction did not depend on the size of imaged neural population, with the same reconstruction quality observed for the same amount of data for $N = 50 - 200$ neurons. In all cases, good reconstructions were obtained with $T \sim 5$–30 min of calcium imaging data.

3.5 Accuracy of the estimates and Fisher information matrix

XXX THIS SHOULD PROBABLY BE REDUCED A BIT XXX

To determine the necessary amount of data for accurate estimation of the functional connectivity matrix, we calculate Fisher information for $P[\mathbf{w}|\mathbf{X}]$. For clarity, we assume here that $\Delta \to 0$, and so $f(J) \approx e^J \Delta$, and that spike trains are known perfectly, thus there is no corruption due to inference from low-SNR calcium imaging data. We also assume that spikes couple only over a single time bin. Then, we write the Fisher information matrix as:

$$C^{-1} = \left[\frac{\partial(-\ln P)}{\partial w_{ij} \partial w_{i'j'}}\right] = - \quad \delta_{ii'} \sum_t \Big[n_i(t) n_j(t-1) n_{j'}(t-1) \left(-\frac{f'(J_i(t))^2}{f(J_i(t))^2} + \frac{f''(J_i(t))}{f(J_i(t))} \right) \\ -(1 - n_i(t)) n_j(t-1) n_{j'}(t-1) f''(J_i(t)) \Big] . \tag{18}$$

where f' and f'' correspond to the first and the second derivatives of our linking function (c.f Eq. 1), and $\delta_{ii'}$ is the Kronecker's delta symbol ($\delta_{ii'} = 1$ for $i = i'$, and $\delta_{ii'} = 0$ otherwise). Letting $f(J) = e^J \Delta$, this may be rewritten as:

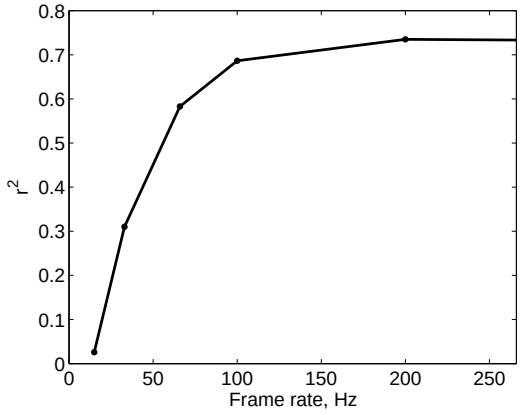

Figure 6: Accuracy of the inferred connectivity weights as function of the frame rate of calcium imaging. Connectivity matrix here was inferred from the original spike trains observed at corresponding frame rates, thus establishing the upper performance bound for inference using calcium imaging data. A network of $N = 25$ neurons, firing at ≈ 5 Hz and imaged for $T = 600$ sec was used.

$$\begin{aligned} C^{-1} &= \delta_{ii'} TP[n_i(t) = 0, n_j(t-\Delta) = 1, n_{j'}(t-\Delta) = 1] E[e^{J_i(t)} | n_i(t) = 0, n_j(t-\Delta) = 1, n_{j'}(t-\Delta) = 1] \\ &= T\left[(r\tau_w)\delta_{ii'}\delta_{jj'} + O((r\tau_w)^2)\right] r. \end{aligned} \tag{19}$$

where τ_w is "the coincidence time" — the typical EPSP/IPSP time-scale over which the spike of one neuron affects the spike probability of the other neuron — and $r \approx E[e^{J_i(t)} | n_i(t) = 0, n_j(t-\Delta) = 1, n_{j'}(t-\Delta) = 1]$ is the typical firing rate. As can be seen from Eq. (19), Fisher information matrix is nearly diagonal, and thus the covariance matrix C, the inverse of the Fisher information, can be computed trivially using:

$$C = \left[(T\Delta)(r^2\tau_w)I + O((r\tau_w)^2)\right]^{-1} = (T\Delta r^2\tau_w)^{-1} I + O((r\tau_w)^2) \tag{20}$$

For successful determination of the functional connectivity matrix $\mathbf{w}$, the variance C should be made smaller than the typical scale $\langle \mathbf{w}^2 \rangle$, i.e.

$$(T\Delta) \sim (\langle \mathbf{w}^2 \rangle r^2 \tau_w)^{-1}. \tag{21}$$

For typical values of $\mathbf{w}^2 \approx 0.1$, $r \approx 5$ Hz and $\tau_w \approx 10$ msec, with this order of magnitude estimate we obtain T of the order of hundred seconds. This theoretical estimate of the necessary amount of fluorescent data is in good agreement with our simulations.

Note also that necessary recording time does not depend on the number of neurons in the imaged network N. This unexpected result is the direct consequence of the special form of C^{-1} in Eq. 19. In particular, when $r\tau_w << 1$, this matrix is dominated by the diagonal term $(T\Delta)(r^2\tau_w)$, and so the Fisher information matrix is predominantly diagonal with the scale $(r^2\tau_w T\Delta)^{-1}$, independent of the number of neurons N. This theoretical result is also directly confirmed in our simulations.

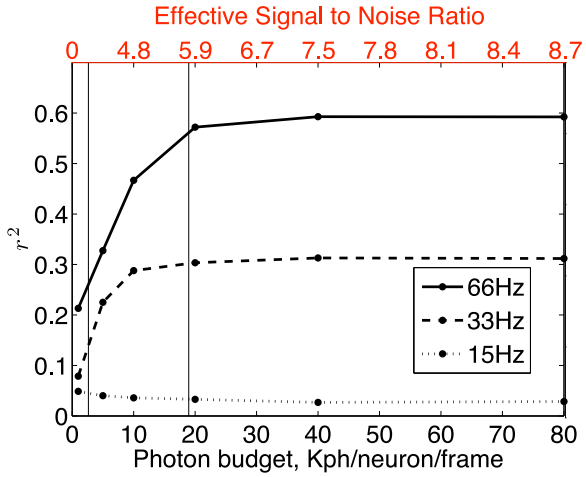

Figure 7: Accuracy of inferred connectivity weights as function of the noise amount in the calcium imaging data, as quantified by experimental photon budget per neuron-frame, for frame rates of 15 Hz, 33 Hz and 66 Hz. Photon counts on the order of 20-40 Kph/frame/neuron are required to achieve the upper bound due by the frame rate. Connectivity matrix here was inferred from simulated fluorescence data using factorized approximation algorithm. Simulation conditions are the same as in Figure 6. Vertical black lines indicate noise levels for the two example traces shown in Figure 3, as determined using the SMC approach to estimate θ_i as described in Section 2.3.

3.6 Impact of using priors on the inference

Taking into account simple prior information about the connectivity matrix resulted in dramatic improvement of the inferred connectivity matrix, Figure 9 and 10. Sparseness prior resulted in dramatic improvements allowing successful reconstruction from as little as 5 min of calcium imaging data, and allowing to achieve for $T \approx 10$ min the same level of accuracy that would otherwise require up to $T \approx 1$ hour of calcium imaging data, Figure 8 and 9. Note, however, that sparse prior resulted in added scale bias into obtained connectivity estimate, thus, effectively destroying information about the scale of connection weights in a population. Furthermore, information about connected neural pairs and about inhibitory or excitatory nature of particular neuron could be reliably obtained, Figure 10. Dale's prior, on the other hand, only led to 10% in the correlation coefficient r^2 of the reconstructed connectivity matrix, and was not found significant.

3.7 Impact of strong correlations and deviations from generative model on the inference

"Anatomical" connectivity was recovered in our experiments despite potential problems noted in the literature [XXX], e.g. such as common input from correlated neurons. This is primarily due to the particular form of the activity in our neural networks, whereas firing of neurons occurred independently, thus, allowing GLM explore the full range of possible input configurations and disentangle common inputs.

Estimation of the functional connectivity is fundamentally routed in observing changes in the spike rate conditioned on the state of the other neurons. Intuitively, such estimation

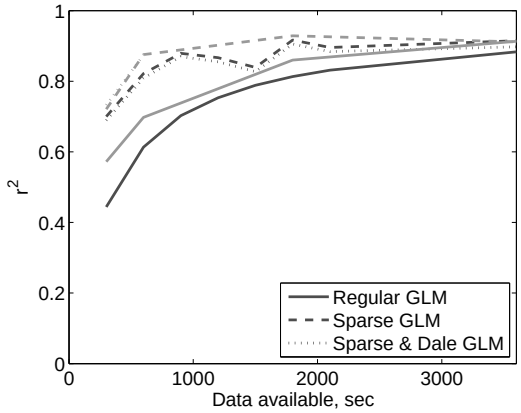

Figure 8: Accuracy of inferred connectivity weights as function of the imaging time and neural population size. Incorporating simple priors such as exponential prior on the connectivity weights allows to boost reconstruction accuracy dramatically (dashed lines). In this latter case, $T = 300$ sec is already sufficient to recover 70% of the variance in the connection weights. Incorporating Dale's prior leads to only marginal improvement (dotted line). As shown in the methods, reconstruction accuracy does not depend on the neural population size N. Here, neural population from $N = 10$ to $N = 200$ were simulated for different T, where $N = 200$ (gray) and $N = 100$ (black) are shown. All networks were prepared in similar state by adjusting strength of inhibitory connections to achieve similar mean firing rate ≈ 5 Hz, although actual firing rate in these networks could vary. In all cases, $T = 5$ min - 0.5 hour is sufficient to produce accurate reconstructions.

can be compared to observing changes in $p(\mathbf{n}(t)) = \exp(\sum_j w_{ij} n_j(t))$ for different neural configurations $\mathbf{n}(t)$, i.e. estimating a vector $\mathbf{w}_i$ from a number of dot-products $\mathbf{w}_i \cdot \mathbf{n}(t)$ with different vectors $\mathbf{n}(t)$. In order to properly estimate all components of $\mathbf{w}_i$ the set of available $\mathbf{n}(t)$ should be rich enough to span all N dimensions of $\mathbf{w}_i$. In case of independent firing such condition is clearly satisfied. Should this condition be violated, however, e.g. due to high correlation between spiking of few neurons, spike trains may not provide access to the complete vector $\mathbf{w}_i$, and the connection weights inferred from such activity data may effectively "aggregate" true connection weights in arbitrary linear combinations.

We carried out a simulation of hypothetical "strongly" coupled neural network, where in addition to weak sparse connectivity we introduced sparse random strong connectivity component. In some sense, we allowed a fraction of neurons to couple strongly to the other neurons, thus making them "command" neurons "driving" activity of the other neurons. The strength of strong connectivity component was chosen to build up the actual firing rate dynamically from the baseline rate of $r = \exp(b) \approx 1$ Hz to ≈ 5 Hz. Such neural network showed patterns of activity very different from the weakly coupled networks inspected above, Figure 11. In particular, large number of highly correlated, synchronously locked firings of many neurons were evident in this network. Likewise, our algorithm was not able to identify the true connectivity matrix correctly, Figure 11.

On the other hand, our inference algorithm showed significant robustness to different deviations from the generative model. One important deviation that is likely to be present in the real data is variations in the time-scales of EPSPs of different synapses. Up to now,

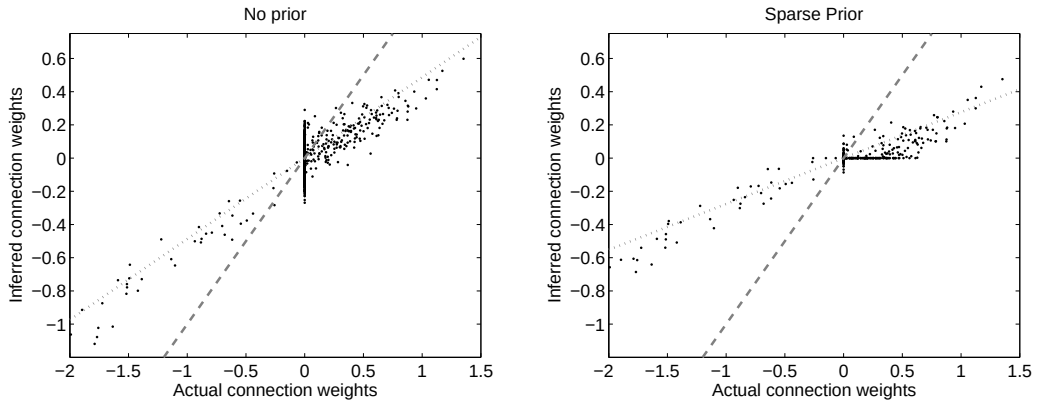

Figure 9: Incorporating simple priors on the distribution of connectivity weights in the Bayesian inference algorithm, such as exponential sparseness prior, is essential to achieve much more accurate reconstructions than using simple GLM from a smaller amount of calcium imaging data. Here, connection weights reconstructed using simple GLM (left panel) or sparse-prior GLM (right panel) are shown in a scatter plot for a network of $N = 50$ neurons firing at ≈ 5 Hz and imaged for $T = 600$ sec. $r^2 = 0.64$ for simple GLM solution and $r^2 = 0.85$ for sparse-GLM solution.

all EPSP time-scales τ_w were assumed to be the same in our inference algorithm. Variability in τ_w would result in added variance in the estimated weights w_{ij} through τ_w dependence of the scaling factor Eq.(17). Still, we found that such added variance to be insignificant in our simulations with τ_w varying for up to 25%, Figure 12.

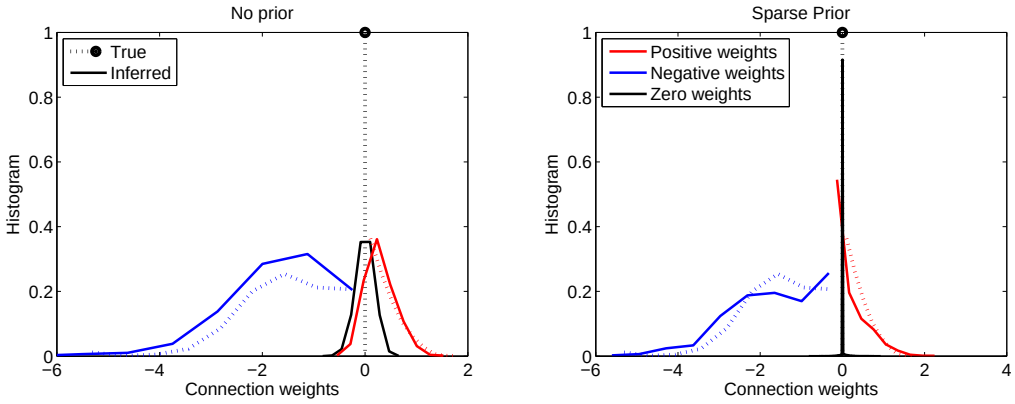

Figure 10: Distribution of inferred connection weights using simple GLM (left) and sparse GLM (right) vs true distributions. When sparse exponential prior on the distribution of connection weights is enacted, dispersion in inferred connection weights is substantially reduced and, in particular, it becomes possible to reliably determine which neural pairs are connected. Distributions are shown for a network of $N = 200$ neurons firing at ≈ 5 Hz and imaged for $T = 600$ s was used here.

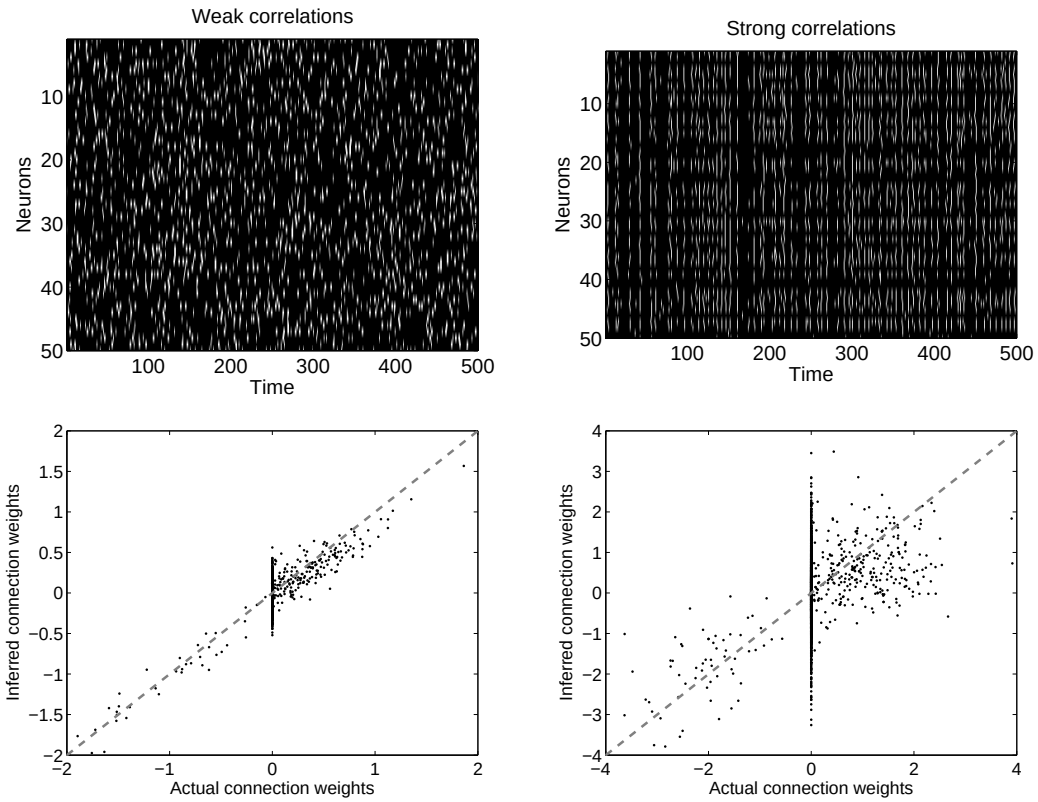

Figure 11: Diverseness of observed neural activity patterns is required for functional connectivity to give access to the actual "anatomical" structure of the neural circuit. Here, 15 sec of simulated spike trains for a weakly coupled network (upper-left) and a network with strongly coupled component (upper-right) are shown. In weakly coupled network spikes are sufficiently uncorrelated to give access to all different neural activity patterns needed to properly estimate true weights $\mathbf{w}_i$. In strongly coupled case, many instances of highly synchronous locked firings are evident, thus preventing observation of sufficiently rich ensemble of activity patterns. Accordingly, GLM solution for the strongly coupled neural network (lower-right) does not represent the true connectivity of the circuit, even for the weakly coupled circuit's component. This is contrary to the weakly-coupled network (lower-left) where true connectivity is successfully estimated. Networks of $N = 50$ neurons firing at ≈ 5 Hz and imaged for $T = 600$ sec were used to produce this figure.

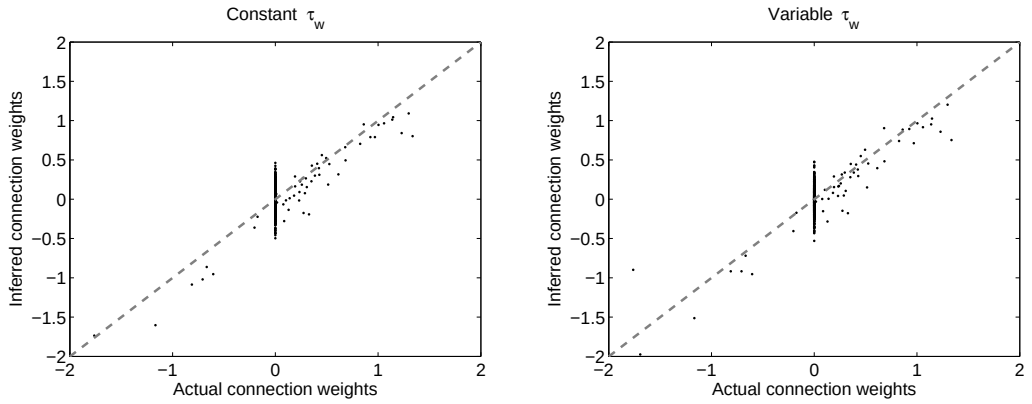

Figure 12: Bayesian inference algorithm is robust to distortions of the underlying generative model. One distortion that should be expected is variability of the EPSP time courses from neuron to neuron, and possibly synapse to synapse. With up to 25% variability allowed in EPSP time scales τ_w (right panel) our algorithm provided reconstructions of the same quality as when all τ_w were the same (left panel). Simulation conditions are the same as in Figure 6.

4 Discussion

The main result of this work is that given only noisy, temporally subsampled, calcium imaging observations from a population of neurons, we can efficiently infer the most likely functional connection matrix. More specifically, we show that for reasonable assumptions about the parameters of the neurons (c.f Figure 7), and about 10 minutes of simulated data (c.f. Figure 8), we can accurately reconstruct the connection matrix, in only about 10 of computational time, when running our algorithm on a high-perform cluster. And while our estimates exhibit some scale error (c.f. Figure 4), the scale error is explained by the poor temporal resolution inherent in the data (c.f. Figure 5). These results depend on a few theoretical advancements. First, we develop an embedded-chain-within-blockwise-Gibbs algorithm for jointly sampling spike trains, given the calcium traces. Then, we show that by factorizing, we can obtain approximately equally good estimates (c.f. Figure 4), which greatly expedites the computations. Finally, we can impose a sparse prior on the connection matrix, justified by recent experimental findings (**?**), to reduce the amount of observations necessary to obtain good estimates (c.f. Figure 9 and 10). And while our approach breaks down in the face of strong correlations (c.f. Figure 11), the algorithm fairs well under certain model misspecifications (c.f. Figure 12).

Importantly, the above-described approach for learning a functional connection matrix differs significantly from previous work. Naively, one can define functional connectivity between a pair of neurons as some function of their two spike trains, $f(\mathbf{n}_1, \mathbf{n}_2)$. Examples of such functions include correlation coefficient, lagged cross-correlation, transfer entropy, and Granger causality. These measures are problematic, in that they implicitly ignore the spike trains of all other neurons. One would rather have a measure of connectivity that is conditioned on the spike trains of all observable neurons, i.e., $f(\mathbf{n}_1, \mathbf{n}_2 | \mathbf{n}_3, \ldots, \mathbf{n}_N)$. Several groups have recently developed methods to infer pairwise connectivity conditioned on all other observable spike trains (**?**; **?**; **?**; **?**; **?**). To our knowledge, however, this work represents the first attempt to infer connectivity given some data other than spike trains. As the spikes contain all the information about connectivity, we therefore must infer the spike trains first, and then estimate the functional connectivity.

While this strategy captures the pairwise dependency between neurons, conditioned on all the other observable spike trains, it is often the case that many unobservable neurons are impacting the spike train statistics. Developing methods to cope with these unobserved neurons is currently an area of active investigation (**?**; **?**). Incorporating these features into our model to further improve our estimates is an important direction for future work.

Along with coping with unobserved neurons, we have considered several other potential directions. First, incorporating photo-stimulation to activate or deactivate small subpopulations of neurons to reduce the uncertainty of our estimates could potentially significantly reduce the amount of experimental time required to obtain a particular r^2 value. Second, a fully Bayesian algorithm, estimating distributions of all the parameters (necessitating establishing priors on each), as opposed to only finding the maximum a posteriori (MAP) estimate, would be desirable. Third, certain adaptations of this approach might be necessary to apply these algorithms to real data, including dealing with multiple spikes per frame, non-stationarities (e.g., bleaching), etc. All these advances are currently being pursued.

Acknowledgements

The authors would like to acknowledge Rafael Yuste, Brendon O. Watson, Adam Packer, Tanya Sippy, Tom Mrsic-Flogel and Vincent Bonin for data and helpful discussions.

References

Abeles, M. (1991). *Corticonics*. Cambridge University Press.

Andrieu, C., Doucet, A., and Holenstein, A. (2007). Particle markov chain monte carlo. *Working paper*.

Bickel, P., Li, B., and Bengtsson, T. (2008). Sharp failure rates for the bootstrap particle filter in high dimensions. In Clarke, B. and Ghosal, S., editors, *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, pages 318–329. IMS.

Binzegger, T., Douglas, R. J., and Martin, K. A. C. (2004). A Quantitative Map of the Circuit of Cat Primary Visual Cortex. *J. Neurosci.*, 24(39):8441–8453.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Oxford University Press.

Braitenberg, V. and Schuz, A. (1998). *Cortex: statistics and geometry of neuronal connectivity.* Springer, Berlin.

Brenowitz, S. D. and Regehr, W. G. (2007). Reliability and heterogeneity of calcium signaling at single presynaptic boutons of cerebellar granule cells. *J Neurosci*, 27(30):7888–7898.

Briggman, K. L. and Denk, W. (2006). Towards neural circuit reconstruction with volume electron microscopy techniques. *Current Opinions in Neurobiology*, 16:562.

Brillinger, D. (1988). Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cyberkinetics*, 59:189–200.

Brillinger, D. (1992). Nerve cell spike train data analysis: a progression of technique. *Journal of the American Statistical Association*, 87:260–271.

Buhl, E., Halasy, K., and Somogyi, P. (1994). Diverse sources of hippocampal unitary inhibitory postynaptic potentials and the number of synaptic release sites. *Nature*, 368:823–828.

Candes, E. J. and Romberg, J. (2005). Practical signal recovery from random projections.

Chornoboy, E., Schramm, L., and Karr, A. (1988). Maximum likelihood identification of neural point process systems. *Biological Cybernetics*, 59:265–275.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal Royal Stat. Soc., Series B*, 39:1–38.

Djurisic, M., Antic, S., Chen, W. R., and Zecevic, D. (2004). Voltage imaging from dendrites of mitral cells: EPSP attenuation and spike trigger zones. *J. Neurosci.*, 24(30):6703–6714.

Donoho, D. and Elad, M. (2003). Optimally sparse representation in general (nonorthogonal) dictionaries via L^1 minimization. *PNAS*, 100:2197–2202.

Doucet, A., de Freitas, N., and N., G. (2001). *Sequential Monte Carlo in Practice*. Springer.

Doucet, A., Godsill, S., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10:197–208.

Escola, S. and Paninski, L. (2008). Hidden Markov models applied toward the inference of neural states and the improved estimation of linear receptive fields. *Under review, Neural Computation*.

Feldmeyer, D., Egger, V., Lubke, J., and Sakmann, B. (1999). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single "barrel" of developing rat somatosensory cortex. *J Physiol*, 521 Pt 1:169–90.

Feldmeyer, D. and Sakmann, B. (2000). Synaptic efficacy and reliability of excitatory connections between the principal neurones of the input (layer 4) and output layer (layer 5) of the neocortex. *J Physiol*, 525:31–9.

Godsill, S., Doucet, A., and West, M. (2004). Monte Carlo smoothing for non-linear time series. *Journal of the American Statistical Association*, 99:156–168.

Gomez-Urquijo, S. M., Reblet, C., Bueno-Lopez, J. L., and Gutierrez-Ibarluzea, I. (2000). Gabaergic neurons in the rabbit visual cortex: percentage, distribution and cortical projections. *Brain Res*, 862:171–9.

Greenberg, D. S., Houweling, A. R., and Kerr, J. N. D. (2008). Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nat Neurosci*.

Gupta, A., Wang, Y., and Markram, H. (2000). Organizing principles for a diversity of gabaergic interneurons and synapses in the neocortex. *Science*, 287:273–8.

Harris, K., Csicsvari, J., Hirase, H., Dragoi, G., and Buzsaki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, 424:552–556.

Hatsopoulos, N., Ojakangas, C., Paninski, L., and Donoghue, J. (1998). Information about movement direction obtained by synchronous activity of motor cortical neurons. *PNAS*, 95:15706–15711.

Helmchen, F., Imoto, K., and Sakmann, B. (1996). Ca^{2+} buffering and action potential-evoked Ca^{2+} signaling in dendrites of pyramidal neurons. *Biophys J*, 70(2):1069–1081.

Iyer, V., Hoogland, T. M., and Saggau, P. (2006). Fast functional imaging of single neurons using random-access multiphoton (RAMP) microscopy. *J Neurophysiol*, 95(1):535–545.

Koch, C. (1999). *Biophysics of Computation*. Oxford University Press.

Kulkarni, J. and Paninski, L. (2007). Common-input models for multiple neural spike-train data. *Network: Computation in Neural Systems*, 18:375–407.

Lefort, S., Tomm, C., Floyd Sarria, J.-C., and Petersen, C. C. H. (2009). The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex. *Neuron*, 61:301–16.

Li, K. and Duan, N. (1989). Regression analysis under link violation. *Annals of Statistics*, 17:1009–1052.

Litke, A., Bezayiff, N., Chichilnisky, E., Cunningham, W., Dabrowski, W., Grillo, A., Grivich, M., Grybos, P., Hottowy, P., Kachiguine, S., Kalmar, R., Mathieson, K., Petrusca, D., Rahman, M., and Sher, A. (2004). What does the eye tell the brain? development of a system for the large scale recording of retinal output activity. *IEEE Trans Nucl Sci*, pages 1434–1440.

Livet, J., Weissman, T., Kang, H., Draft, R., Lu, J., Bennis, R., Sanes, J., and Lichtman, J. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, 450:56–62.

Luczak, A., Bartho, P., Marguet, S., Buzsaki, G., and Harris, K. (2007). Sequential structure of neocortical spontaneous activity in vivo. *PNAS*, 104:347–352.

McLachlan, G. and Krishnan, T. (1996). *The EM Algorithm and Extensions*. Wiley-Interscience.

Micheva, K. and Smith, S. (2007). Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron*, 55:25–36.

Mishchenko, Y. (2009). Strategies for identifying exact structure of neural circuits with broad light microscopy connectivity probes. *Preprint: http://precedings.nature.com/documents/2669/version/2*.

Mishchenko, Y., Spacek, J., Mendenhall, J., Chklovskii, D., and Harris, K. M. (2009). Reconstruction of hippocampal ca1 neuropil at nanometer resolution reveals disordered packing of processes and dependence of synaptic connectivity on local environment and dendritic caliber. *Preprint: http://precedings.nature.com/documents/2669/version/2*.

Neal, R., Beal, M., and Roweis, S. (2003). Inferring state sequences for non-linear systems with embedded hidden Markov models. *NIPS*, 16.

Ng, A. (2004). Feature selection, L_1 vs. L_2 regularization, and rotational invariance. *ICML*, 21.

Nguyen, Q. T., Callamaras, N., Hsieh, C., and Parker, I. (2001). Construction of a two-photon microscope for video-rate Ca^{2+} imaging. *Cell Calcium*, 30(6):383–393.

Nykamp, D. (2007). A mathematical framework for inferring connectivity in probabilistic neuronal networks. *Mathematical Biosciences*, 205:204–251.

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, 15:243–262.

Paninski, L., Fellows, M., Shoham, S., Hatsopoulos, N., and Donoghue, J. (2004). Superlinear population encoding of dynamic hand trajectory in primary motor cortex. *J. Neurosci.*, 24:8551–8561.

Petersen, C. C. and Sakmann, B. (2000). The excitatory neuronal network of rat layer 4 barrel cortex. *J Neurosci*, 20(20):7579–86.

Petrusca, D., Grivich, M. I., Sher, A., Field, G. D., Gauthier, J. L., Greschner, M., Shlens, J., Chichilnisky, E. J., and Litke, A. M. (2007). Identification and characterization of a Y-like primate retinal ganglion cell type. *J. Neurosci.*, 27(41):11019–11027.

Pillow, J., Shlens, J., Paninski, L., Sher, A., Litke, A., Chichilnisky, E., and Simoncelli, E. (2008). Spatiotemporal correlations and visual signaling in a complete neuronal population. *Nature*, 454:995–999.

Plesser, H. and Gerstner, W. (2000). Noise in integrate-and-fire neurons: From stochastic input to escape rates. *Neural Computation*, 12:367–384.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77:257–286.

Ramon y Cajal, S. (1904). *La Textura del Sistema Nerviosa del Hombre y los Vertebrados*. Moya.

Ramon y Cajal, S. (1923). *Recuerdos de mi vida: Historia de mi labor cientifica*. Alianza Editorial.

Reddy, G., Kelleher, K., Fink, R., and Saggau, P. (2008). Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity. *Nature Neuroscience*, 11(6):713–720.

Reddy, G. D. and Saggau, P. (2005). Fast three-dimensional laser scanning scheme using acousto-optic deflectors. *J Biomed Opt*, 10(6):064038.

Reyes, A., Lujan, R., Rozov, A., Burnashev, N., Somogyi, P., and Sakmann, B. (1998). Target-cell-specific facilitation and depression in neocortical circuits. *Nat Neurosci*, 1:279–285.

Rigat, F., de Gunst, M., and van Pelt, J. (2006). Bayesian modelling and analysis of spatio-temporal neuronal networks. *Bayesian Analysis*, 1:733–764.

Robert, C. and Casella, G. (2005). *Monte Carlo Statistical Methods*. Springer.

Salome, R., Kremer, Y., Dieudonne, S., Leger, J.-F., Krichevsky, O., Wyart, C., Chatenay, D., and Bourdieu, L. (2006). Ultrafast random-access scanning in two-photon microscopy using acousto-optic deflectors. *Journal of Neuroscience Methods*, 154(1-2):161–174.

Santhanam, G., Ryu, S. I., Yu, B. M., Afshar, A., and Shenoy, K. V. (2006). A high-performance brain-computer interface. *Nature*, 442:195–198.

Sayer, R. J., Friedlander, M. J., and Redman, S. J. (1990). The time course and amplitude of epsps evoked at synapses between pairs of ca3/ca1 neurons in the hippocampal slice. *J. Neurosci.*, 10:826–36.

Segev, R., Goodhouse, J., Puchalla, J., and Berry, M. (2004). Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nature Neuroscience*, 7:1154–1161.

Shumway, R. and Stoffer, D. (2006). *Time Series Analysis and Its Applications*. Springer.

Song, S., Sjostrom, P. J., Reiql, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biology*, 3:e68.

Stein, R. B., Weber, D. J., Aoyagi, Y., Prochazka, A., Wagenaar, J. B. M., Shoham, S., and Normann, R. A. (2004). Coding of position by simultaneously recorded sensory neurones in the cat dorsal root ganglion. *J Physiol (Lond)*, 560(3):883–896.

Stevenson, I., Rebesco, J., Hatsopoulos, N., Haga, Z., Miller, L., and Koerding, K. (2008). Inferring network structure from spikes. *Statistical Analysis of Neural Data meeting*.

Stevenson, I. H., Rebesco, J. M., Hatsopoulos, N. G., Haga, Z., Miller, L. E., and Kording, K. P. (2009). Bayesian inference of functional connectivity and network structure from spikes. *IEEE Trans. Neural Systems and Rehab.*, 17:203–13.

Stosiek, C., Garaschuk, O., Holthoff, K., and Konnerth, A. (2003). In vivo two-photon calcium imaging of neuronal networks. *Proceedings of The National Academy Of Sciences Of The United States Of America*, 100(12):7319–7324.

Thompson, A., Girdlestone, D., and West, D. (1988). Voltage-dependent currents prolong single-axon postsynaptic potentials in layer III pyramidal neurons in rat neocortical slices. *J Neurophysiol*, 60:1896–1907.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58:267–288.

Tipping, M. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244.

Truccolo, W., Eden, U., Fellows, M., Donoghue, J., and Brown, E. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble and extrinsic covariate effects. *Journal of Neurophysiology*, 93:1074–1089.

Tsien, R. Y. (1989). Fluorescent probes of cell signaling. *Ann. Rev. Neurosci.*, 12:227–253.

Vidne, M., Kulkarni, J., Ahmadian, Y., Pillow, J., Shlens, J., Chichilnisky, E., Simoncelli, E., and Paninski, L. (2009). Inferring functional connectivity in an ensemble of retinal ganglion cells sharing a common input. *COSYNE*.

Vogelstein, J. T., Watson, B. O., Packer, A. M., Yuste, R., Jedynak, B., and Paninski, L. (2009). Spike inference from calcium imaging using sequential monte carlo methods. *Preprint*.

Wallace, D., zum Alten Borgloh, S., Astori, S., Yang, Y., Bausen, M., K
"ugler, S., Palmer, A., Tsien, R., Sprengel, R., Kerr, J., Denk, W., and Hasan, M. (2008). Single-spike detection in vitro and in vivo with a genetic Ca2+ sensor. *Nature methods*, 5(9):797–804.

Yaksi, E. and Friedrich, R. W. (2006). Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca^{2+} imaging. *Nature Methods*, 3(5):377–383.

Yasuda, R., Nimchinsky, E. A., Scheuss, V., Pologruto, T. A., Oertner, T. G., Sabatini, B. L., and Svoboda, K. (2004). Imaging calcium concentration dynamics in small neuronal compartments. *Sci STKE*, 219:p15.

Yuste, R., Konnerth, A., Masters, B., et al. (2006). *Imaging in Neuroscience and Development, A Laboratory Manual*.